SIAM J. CONTROL AND OPTIMIZATION
Vol. 17, No. 3, May 1979

0363-0129/79/1703-0003 $01.00/0

ON THE STOCHASTIC REALIZATION PROBLEM*

ANDERS LINDQUIST† AND GIORGIO PICCI‡

Abstract. Given a mean square continuous stochastic vector process y with stationary increments and a rational spectral density Φ such that $\Phi(\infty)$ is finite and nonsingular, consider the problem of finding all minimal (wide sense) Markov representations (*stochastic realizations*) of y. All such realizations are characterized and classified with respect to deterministic as well as probabilistic properties. It is shown that only certain realizations (*internal stochastic realizations*) can be determined from the given output process y. All others (*external stochastic realizations*) require that the probability space be extended with an exogeneous random component. A complete characterization of the sets of internal and external stochastic realizations is provided. It is shown that the state process of any internal stochastic realization can be expressed in terms of two steady-state Kalman–Bucy filters, one evolving forward in time over the infinite past and one backward over the infinite future. An algorithm is presented which generates families of external realizations defined on the same probability space and totally ordered with respect to state covariances.

1. Introduction. One of the most common models of random phenomena in control theory is provided by the linear stochastic system

$$dx = Ax\,dx + B\,dw, \tag{1.1a}$$

$$dz = Cx\,dt + D\,dw, \tag{1.1b}$$

where A, B, C and D are constant matrices of dimensions $n\times n$, $n\times k$, $m\times n$ and $m\times k$ respectively, and w is a k-dimensional mean-square continuous stochastic process with zero mean, stationary orthogonal increments, and $w(0)=0$. Here we shall assume that w is defined on the whole real line R, that is

$$E\{w(t)\}=0 \quad \text{for all } t\in R, \qquad E\{w(t)w(s)'\}=\tfrac{1}{2}\{|t|+|s|-|t-s|\}I \tag{1.2}$$

[35; p. 51], where $E\{\cdot\}$ denotes mathematical expectation and prime ($'$) transposition. (All vectors without prime are column vectors.) For later reference, let $\mathcal{W}_k$ denote the class of all such orthogonal increment processes, the index referring to the dimension; more generally we shall say that the process is of class $\mathcal{W}$. Moreover, we assume that A is a stability matrix, i.e. all the eigenvalues of A are situated in the left complex half-plane; we shall write $\operatorname{Re}\{\lambda(A)\}<0$ for short. This assumption will insure that (1.1a) has the unique solution

$$x(t)=\int_{-\infty}^{t} e^{A(t-\tau)}B\,dw(\tau) \tag{1.3}$$

on the real line, where the integral is defined in quadratic mean. This is an n-dimensional vector process. If, in addition, we assume that $z(0)=0$, the m-dimensional process z can be determined uniquely by integrating (1.1b). We shall call x the *state process*, w the *input process* and z the *output process*. Clearly the state process x is (wide sense) stationary, i.e. the *state covariance matrix*

$$P=E\{x(t)x(t)'\} \tag{1.4}$$

* Received by the editors June 14, 1977 and in revised form March 15, 1978.

† Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506, U.S.A. This work was supported by the National Science Foundation under Grant MPS 75-07028.

‡ Laboratorio per Ricerche di Dinamica dei Sistemi e di Elettronica Biomedica, Consiglio Nazionale delle Ricerche, Casella Postale 1075, 35100 Padova, Italy. This work was supported partially by the National Science Foundation under Grant MPS 75-07028 and partially by the Air Force Office of Scientific Research under Grant AFOSR-78-3519.

does not depend on t, and it satisfies the Lyapunov equation

$$AP+PA'+BB'=0. \tag{1.5}$$

(See e.g. [35].) The output process z has stationary increments.

Each $w \in \mathcal{W}_k$ has a unique spectral representation

$$w(t)=\int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega}\, d\hat{w}(\omega) \tag{1.6}$$

[12; p. 205], where $d\hat{w}$ is an orthogonal stochastic measure such that $E\{d\hat{w}(\omega)\, d\hat{w}(\omega)\dagger\}=I\, d\omega$. (Here $\dagger$ denotes the complex conjugation and transposition.) Then (1.3) may be written

$$x(t)=\int_{-\infty}^{\infty} e^{i\omega t}(i\omega I-A)^{-1}B\, d\hat{w}(\omega). \tag{1.7a}$$

(Indeed, making the substitution $(sI-A)^{-1}=(1/s)[I+A(sI-A)^{-1}]$, (1.7a) is seen to satisfy (1.1a.) Inserting (1.7a) into (1.1b) and integrating yields

$$z(t)=\int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega} W(i\omega)\, d\hat{w}(\omega), \tag{1.7b}$$

where

$$W(s)=C(sI-A)^{-1}B+D. \tag{1.8}$$

We shall call W the *transfer function* of (1.1). Relation (1.7b) is a spectral representation of z; $d\hat{z}(\omega):=W(i\omega)\, d\hat{\omega}(\omega)$ being an orthogonal stochastic measure such that

$$E\{d\hat{z}(\omega)\, d\hat{z}(\omega)\dagger\}=\Phi(i\omega)\, d\omega, \tag{1.9}$$

where Φ is the *spectral density* given by

$$\Phi(s)=W(s)W(-s)'. \tag{1.10}$$

This is an $m\times m$-matrix of rational functions such that (i) each element of Φ is analytic on the imaginary axis, (ii) Φ is parahermitian, i.e. $\Phi(-s)=\Phi(s)'$, (iii) $\Phi(i\omega)$ is nonnegative definite Hermitian for all real ω, and (iv) $\Phi(\infty)<\infty$. Such a Φ is called a *spectral function* [3], [4].

In this paper we consider the following inverse problem. Let $\{y(t);\, t\in R\}$ be a given mean-square continuous and purely nondeterministic m-dimensional stochastic process with zero mean, stationary increments and $y(0)=0$. Then there is a spectral representation

$$y(t)=\int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega}\, d\hat{y}(\omega) \tag{1.11}$$

[12; p. 205], where $d\hat{y}$ is an orthogonal stochastic measure such that [9]

$$E\{d\hat{y}(\omega)\, d\hat{y}(\omega)\dagger\}=\Phi(i\omega)\, d\omega. \tag{1.12}$$

Here Φ is an $m\times m$-matrix of real rational functions satisfying conditions (i)–(iv) above. Setting $R:=\Phi(\infty)$, we also assume that (v) R^{-1} exists and that (vi) $\Phi(i\omega)$ is positive definite for all real ω. The problem is to find representations (1.1) such that the output process z is equivalent to the given process y in some sense to be specified below. Such a representation will be called a *stochastic realization*.

More precisely, the system (1.1) will be called a *wide sense stochastic realization* of y if z has the same spectral density Φ as y and a *proper stochastic realization* if, for each

$t \in (-\infty, \infty)$, $z(t) = y(t)$, a.s. (In the sequel we shall leave out the "a.s.", hence regarding such equivalent processes as equal.) Clearly each proper stochastic realization is also a wide sense stochastic realization, but the converse is not true.

The stochastic realization problem is related to the *spectral factorization problem*: Given a rational spectral function Φ, find all matrices $W(s)$ of real rational functions with all its poles in $\operatorname{Re}(s) < 0$ and satisfying (1.10). Such a function will be called a *stable spectral factor*. Let $\delta\{\cdot\}$ denote McMillan degree [8]. Then $\delta\{W\} \geqq \frac{1}{2}\delta\{\Phi\}$; if there is equality we shall say that W is *minimal*. We have seen that the transfer function (1.8) of any wide sense stochastic realization of y is a stable spectral factor of the spectral density of y. Conversely any such spectral factor W is the transfer function of an equivalence class of wide sense stochastic realizations. In fact, for *any* orthogonal stochastic measure $d\hat{w}$ such that $E\{d\hat{w}(\omega)\, d\hat{w}(\omega)^\dagger\} = I\, d\omega$, the process

$$z(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega t} - 1}{i\omega} W(i\omega)\, d\hat{w}(\omega) \tag{1.13}$$

has the same spectral density as y. Since W is a real rational matrix function analytic in $\operatorname{Re}(s) \geqq 0$, there is a quadruplet $[A, B, C, D]$ of matrices such that (1.8) holds [8], with A a stability matrix. Now let x be defined by (1.7a) and w by (1.6). Then w is of class $\mathcal{W}$ and (x, z) satisfy (1.1) as asserted. Note that $[A, B, C, D]$ defines one wide sense stochastic realization for *each* $w \in \mathcal{W}_k$. Since these realizations are equivalent up to second-order properties of z, in the sequel we shall say that $[A, B, C, D]$ is a wide sense stochastic realization, thereby referring to the whole equivalence class. To avoid trivialities we shall assume that the representation (1.8) is chosen so that the dimension of the matrix A equals $\delta(W)$, i.e. we shall only consider quadruplets $[A, B, C, D]$ for which (A, B) is controllable and (A, C) is observable [8]. We shall call a stochastic realization *minimal* if it corresponds to a minimal spectral factor. Hence, the minimal stochastic realizations are precisely those representations (1.1) which have a state process of smallest possible dimension, i.e. $n = \frac{1}{2}\delta(\Phi)$. In this paper we shall restrict our attention to such realizations, the basic problem being to find all of them.

Determining all wide sense minimal stochastic realizations $[A, B, C, D]$ is a deterministic problem which has been studied extensively by, among others, B. D. O. Anderson [5], Faurre [11] and J. C. Willems [32], the first of whom has named it the *inverse problem of covariance generation*. To facilitate its solution we note that the spectral density of y can be written

$$\Phi(s) = Z(s) + Z(-s)', \tag{1.14}$$

where Z is positive real[1] and rational, and $\delta(Z) = n$ [3], [4], [11], [32]. Let

$$Z(s) = H(sI - F)^{-1}G + \tfrac{1}{2}R \tag{1.15}$$

be a minimal realization [8] of Z, i.e. F, G and H are constant matrices of dimensions $n \times n$, $n \times m$ and $m \times n$ respectively. Hence F is a stability matrix, (F, G) is controllable and (H, F) is observable [8]. There are computational procedures for determining (F, G, H, R) from Φ [8], [13], [31], [38], so in the sequel we shall assume that such a quadruplet is given.

It can be shown [5] that all wide sense minimal stochastic realizations are given by

$$[A, B, C, D] = [TFT^{-1}, T(B_1, B_2)S, HT^{-1}, (R^{1/2}, 0)S] \tag{1.16}$$

[1] A real rational function Z without poles on the imaginary axis is said to be *positive real* if it has no poles in $\operatorname{Re}[s] > 0$ and $Z(i\omega) + Z(-i\omega)'$ is nonnegative definite Hermitian for all real ω.

where the nonsingular matrix T and the orthogonal matrix S are arbitrary, $R^{1/2}$ is the symmetric square-root of R, and (B_1, B_2) are two matrices, $n \times m$ and $n \times p$ respectively (p is arbitrary), such that (P, B_1, B_2) satisfy the conditions

(1.17a) $$FP+PF'+B_1B_1'+B_2B_2'=0$$

(1.17b) $$PH'+B_1R^{1/2}=G,$$

(1.17c) P is a symmetric, positive definite $n \times n$-matrix.

Conversely, any $[A, B, C, D]$ constructed in this fashion is a wide sense minimal realization. It is no restriction to set $T=I$ and $S=I$ in (1.16), i.e. to consider only realizations of the form

(1.18a) $$dx = Fx\,dt + B_1\,du + B_2\,dv,$$

(1.18b) $$dz = Hx\,dt + R^{1/2}\,du$$

where $w = \binom{u}{v} \in \mathscr{W}_{m+p}$. In fact, all other stochastic realizations can be obtained from (1.18) by multiplying (1.18a) by an arbitrary T and transforming w by an orthogonal transformation. Consequently we shall be working in a fixed coordinate system, thereby identifying each transfer function (spectral factor) W with one quadruplet $[F, B, H, (R^{1/2}, 0)]$. Hence the wide sense problem is reduced to determining $B = (B_1, B_2)$.

The main topic of this paper is the characterization of all *proper* minimal stochastic realizations. This is a probabilistic problem. In addition to the input-output map of (1.1) we need to determine the input process w, which is no longer arbitrary; hence we shall be looking for quintuplets $[A, B, C, D, w]$. For an arbitrary representation (1.1), let $(\Omega, \mathscr{F}, P)$ be a probability space on which both y and w are defined, and define $H(y)$ and $H(w)$ to be the closed linear hulls in $L_2(\Omega, \mathscr{F}, P)$ of $\{y_i(t);\, t \in (-\infty, \infty),\, i=1, 2, \cdots, m\}$ and $\{w_i(t);\, t \in (-\infty, \infty),\, i=1, 2, \cdots, k\}$ respectively. Since y is given, $H(y)$ is fixed, whereas $H(w)$ varies with different choices of representation (1.1). For a proper stochastic realization we will always have $H(y) \subset H(w)$. We shall say that $[A, B, C, D, w]$ is an *internal stochastic realization* if $H(y) = H(w)$ and an *external stochastic realization* if $H(y) \neq H(w)$, adding the attribute *minimal* as appropriate. Hence the internal realizations are precisely those proper stochastic realizations which can be constructed in terms of the given process y, whereas the external realizations require extending our probabilistic setting with an exogeneous noise generator unrelated to y. Various aspects of the proper stochastic realization problem have been studied by Akaike [1], [2], Picci [23], [24] and Rozanov [26], but here we shall give a *complete* characterization of all such realizations. (In [21] the internal realizations are constructed from basic principles without first assuming that they are defined by models of type (1.1).) After submitting this paper we have learned about a series of as yet unpublished papers by Ruckebusch [27]–[29] containing discrete-time counterparts of some of the results presented here; these papers provide an alternative approach to the problem.

The outline of the paper goes as follows. Section 2 is devoted to preliminaries and definitions. In § 3 we show that to each proper stochastic realization there is a representation (1.1) with $\mathrm{Re}\,\{\lambda(A)\} > 0$ and $z = y$, the dynamic relations of which evolve backward in time. These representations, which are an important tool in our subsequent analysis, are called *proper backward stochastic realizations*. In §§ 4 and 5 all internal stochastic realizations are characterized, and it is shown that these are precisely the proper stochastic realizations for which $B_2 = 0$. Each internal state process can be expressed in terms of two steady-state Kalman–Bucy estimates, one filter evolving in the forward direction from time $t=-\infty$ and the other in the backward direction from $t=\infty$. Sections 6 and 7 are devoted to external stochastic realizations. First, in § 6, we construct a system of differential equations in B_1 and B_2 which generates families of wide sense stochastic realizations, totally ordered with respect to state covariances. In § 7 this result is interpreted in terms of proper stochastic realizations and a complete characterization of all such realizations is provided.

This paper extends the results reported (without proofs) in our short note [20].

2. Preliminaries and definitions. Let the function Λ: $R^{n\times n}\to R^{n\times n}$ be given by

$$\Lambda(P)=FP+PF'+(G-PH')R^{-1}(G-PH')', \tag{2.1}$$

and define the set $\mathcal{P}=\{P|P'=P;\ \Lambda(P)\leqq 0\}$ of symmetric $n\times n$-matrices, where $Q\geqq 0$ ($Q>0$) means that Q is nonnegative (positive) definite. Also introduce the subset $\mathcal{P}_0=\{P\in\mathcal{P}|\Lambda(P)=0\}$.

In the following theorem we collect some facts from Anderson [5], Faurre [11] and Willems [32].

THEOREM 2.1. *The set $\mathcal{P}$ is closed, bounded and convex, and there are two elements P_* and P^* in $\mathcal{P}_0$ such that*

$$P_*\leqq P\leqq P^* \quad \text{for all } P\in\mathcal{P}. \tag{2.2}$$

Moreover, $\mathcal{P}$ is the set of all solutions P of (1.17), *and $\mathcal{P}_0$ is the set of all such solutions for which $B_2=0$.*

Each $P\in\mathcal{P}$ can be interpreted as the covariance matrix (1.4) of the corresponding stochastic realization (1.18). Consequently, there is a minimum-variance (P_*) and a maximum-variance (P^*) wide sense stochastic realization, and for these realizations we have $B_2=0$.

For each $P\in\mathcal{P}$, define the *feedback matrix*

$$\Gamma=F-(G-PH')R^{-1}H, \tag{2.3}$$

the significance of which will be made clear below. Let the feedback matrices corresponding to P_* and P^* be denoted Γ_* and Γ^* respectively. It can be shown that $\mathrm{Re}\,\{\lambda(\Gamma_*)\}<0$ and $\mathrm{Re}\,\{\lambda(\Gamma^*)\}>0$ [32, p. 260], [11, p. 53]. Consequently, for each matrix N, the Lyapunov equation

$$\Gamma_*'M+M\Gamma_*+H'R^{-1}H+N=0 \tag{2.4}$$

has a unique solution $M_*(N)$, which is positive definite whenever N is nonnegative definite. In fact, since (F, H) is controllable, so is (Γ, H). (See e.g. [36].) Likewise

$$-\Gamma^{*\prime}M-M\Gamma^*+H'R^{-1}H+N=0 \tag{2.5}$$

has a unique positive definite solution $M^*(N)$ for each $N\geqq 0$. Furthermore, define $\mathcal{P}_+=\{P\in\mathcal{P}|P>P_*\}$ and $\mathcal{P}_-=\{P\in\mathcal{P}|P<P^*\}$. Since $\Phi(i\omega)>0$ for all real ω, $P_*<P^*$ [32, p. 360], and consequently $\mathcal{P}_+$ and $\mathcal{P}_-$ are nonempty.

THEOREM 2.2. *Let Π and $\bar{\Pi}$ be the unique solutions of the $n\times n$-matrix differential equations*

$$\dot{\Pi}(t)=\Lambda(\Pi(t));\qquad \Pi(0)=0 \tag{2.6}$$

and

$$\dot{\bar{\Pi}}(t)=\bar{\Lambda}(\bar{\Pi}(t));\qquad \bar{\Pi}(0)=0 \tag{2.7}$$

respectively, where Λ is given by (2.1) *and $\bar{\Lambda}$ by*

$$\bar{\Lambda}(P)=F'P+PF+(H'-PG)R^{-1}(H'-PG)'. \tag{2.8}$$

Then $\Pi(t)\to P_*$ *and* $\bar{\Pi}(t)\to(P^*)^{-1}$ *as* $t\to\infty$. *Moreover, the matrix* $P=P_*+[M_*(N)]^{-1}$ *belongs to* $\mathcal{P}_+$ *if and only if* $N\geqq 0$. *Likewise,* $P=P^*-[M^*(N)]^{-1}$ *belongs to* $\mathcal{P}_-$ *if and only if* $N\geqq 0$. *Finally,* $P^*-P_*=[M_*(0)]^{-1}=[M^*(0)]^{-1}$.

Various versions of this theorem can be found in [7] and [11]. It provides us with a procedure to determine all elements in $\mathcal{P}_+\cup\mathcal{P}_-$: First compute P_* and P^*. Then varying N over the nonnegative cone will generate the other elements in $\mathcal{P}_+\cup\mathcal{P}_-$. The corresponding wide sense stochastic realizations $[F, B, H, (R^{1/2}, 0)]$ can then be obtained by determining $B=(B_1, B_2)$ from

$$B_1=(G-PH')R^{-1/2}, \tag{2.9a}$$

$$B_2B_2'=-\Lambda(P), \tag{2.9b}$$

which is merely (1.17) reformulated.

In § 6 another method for generating wide sense stochastic realizations is presented, which is formulated directly in terms of B, the unknown quantity in $[F, B, H', (R^{1/2}, 0)]$. Hence the intermediate step of determining P will be eliminated. Define $\mathcal{B}$ to be the set of all $B=(B_1, B_2)$ given by (2.9) as P ranges over $\mathcal{P}$. Let $\mathcal{B}_0$, $\mathcal{B}_+$ and $\mathcal{B}_-$ be defined analogously in terms of $\mathcal{P}_0$, $\mathcal{P}_+$ and $\mathcal{P}_-$. The set $\mathcal{B}_0$ consists of all $B\in\mathcal{B}$ with $B_2=0$ (Theorem 2.1). In particular, let B_* and B^* be the unique elements in $\mathcal{B}_0$ corresponding to P_* and P^* respectively.

All stochastic processes in this paper will have finite second order moments. Given a k-dimensional vector process η of this type, defined on some probability space $(\Omega, \mathcal{F}, P)$, and a subset I of $(-\infty, \infty)$, let $H_I(\eta)$ be the closed linear hull in $L_2(\Omega, \mathcal{F}, P)$ of the stochastic variables $\{\eta_i(t);\, t\in I, i=1, 2, \cdots, k\}$. (We write $H_t(\eta)$ if the set I contains only the point t.) If ξ is an l-dimensional stochastic vector such that $\xi_i\in H_I(\eta)$, $i=1, 2, \cdots, l$, we shall misuse notations slightly by writing $\xi\in H_I(\eta)$. For $\zeta\in L_2(\Omega, \mathcal{F}, P)$, let $\hat{E}\{\zeta|H_I(\eta)\}$ be the projection of ζ onto $H_I(\eta)$, i.e. the wide sense conditional mean in the terminology of Doob [10]. (We shall sometimes write $\hat{E}\{\zeta|\eta(t)\}$ instead of $\hat{E}\{\zeta|H_t(\eta)\}$.) For simplicity let $H(\eta)$, $H_t^-(\eta)$ and $H_t^+(\eta)$ denote $H_{(-\infty,\infty)}(\eta)$, $H_{(-\infty,t)}(\eta)$ and $H_{[t,\infty)}(\eta)$ respectively. Moreover, set $\eta_t(\tau)=\eta(t+\tau)-\eta(t)$, and define $H_t^-(d\eta)$ and $H_t^+(d\eta)$ to be respectively $H_0^-(\eta_t)$ and $H_0^+(\eta_t)$. Note that if $\eta(0)=0$ (which is often the case with the processes studied in this paper), we have $H_\infty^-(d\eta)=H(\eta)$.

As mentioned in § 1, any mean-square continuous stochastic vector process $\{\eta(t);\, t\in R\}$ with stationary increments and $\eta(0)=0$ has a representation of the form

$$\eta(t)=\int_{-\infty}^{\infty}\frac{e^{i\omega t}-1}{i\omega}\,d\hat{\eta}(\omega) \tag{2.10}$$

[12; p. 205], where $d\hat{\eta}$ is an orthogonal stochastic measure, called the *stochastic spectral measure* of η. If, in addition, η is purely nondeterministic, it has an absolutely continuous spectral distribution [9], i.e.

$$E\{d\hat{\eta}(\omega)\, d\hat{\eta}(\omega)^\dagger\}=S(i\omega)\, d\omega \tag{2.11}$$

where S is the *spectral density* of η. If $E\{\eta(t)\}=0$ for all t and $S=I$ (identity), η is said to be of class $\mathcal{W}$. The spectral decomposition (2.10) defines an isometric correspondence between $H(\eta)$ and $L_2(R, S(i\omega)\, d\omega)$ under which $\eta(t)$ corresponds to $(e^{i\omega t}-1)/i\omega$; hence to any real random variable $\xi\in H(\eta)$ there corresponds an (essentially) unique $g\in L_2(R, S(i\omega)\, d\omega)$ such that

$$\xi=\int_{-\infty}^{\infty}g(\omega)\, d\hat{\eta}(\omega).$$

In fact, the system of functions $\{(e^{i\omega t}-1)/i\omega;\, t\in R\}$ is complete in $L_2(R, S(i\omega)\, d\omega)$[12; p. 204]. Hence we have the following lemma which we shall need below.

LEMMA 2.3. *Let ξ and η be mean-square continuous and purely nondeterministic stochastic vector processes, defined on the whole real line R, with (jointly) stationary increments and such that $\xi(t)\in H(\eta)$ for all $t\in R$. Let $S(i\omega)$ be the spectral density of η, and assume that $\xi(0)=0$. Then there is a matrix-valued function K such that $((e^{i\omega t}-1)/i\omega)\cdot K(i\omega)\in L_2(R, S(i\omega)\, d\omega)$ for all $t\in R$ and such that*

$$\xi(t)=\int_{-\infty}^{\infty}\frac{e^{i\omega t}-1}{i\omega}K(i\omega)\, d\hat{\eta}\,(\omega). \tag{2.12}$$

If, in addition, ξ and η are both of class $\mathscr{W}$,

$$K(s)K(-s)'=I. \tag{2.13}$$

The last statement follows from $d\hat{\xi}=K(i\omega)\, d\hat{\eta}$ and the fact that both ξ and η have identity spectral densities.

3. Forward and backward stochastic realizations. Let $\{x(t);\, t\in R\}$ *be an* n-dimensional wide sense Markov process, i.e.

$$\hat{E}\{x(s)|H_t^-(x)\}=\hat{E}\{x(s)|x(t)\}\quad \text{for } s\geqq t, \tag{3.1}$$

or equivalently

$$\hat{E}\{x(s)|H_t^+(x)\}=\hat{E}\{x(s)|x(t)\}\quad \text{for } s\leqq t. \tag{3.2}$$

In addition, assume that x is purely nondeterministic and (wide sense) stationary. It is well-known [11] that such a process can be described as the solution of a system of linear stochastic differential equations of the type

$$dx=Ax\, dt+B\, dw, \tag{3.3}$$

where A and B are constant matrices, $\operatorname{Re}\{\lambda(A)\}<0$, and w is a vector process of class $\mathscr{W}$ such that[2] $H_t^+(dw)\perp H_t^-(x)$ for all $t\in R$. [In fact, A being a stability matrix implies that (3.3) has the solution (1.3), and consequently $H_t^-(x)\subset H_t^-(dw)\perp H_t^+(dw)$.] Moreover, the covariance matrix $P:=E\{x(t)x(t)'\}$ satisfies (1.5). The model (3.3) is clearly unsymmetric with respect to time, $x(t)$ being orthogonal to future increments of w, but not to past ones. Hence we shall call (3.3) the *forward representation* of x.

We shall now show that x has a *backward representation* also, i.e. a model (3.3) with $\operatorname{Re}\{\lambda(A)\}>0$ and $H_t^-(dw)\perp H_t^+(x)$ for all $t\in R$. To this end first observe that the forward representation (3.3) can be integrated between t and s to yield

$$x(s)=e^{A(s-t)}x(t)+\int_t^s e^{A(s-\tau)}B\, dw(\tau), \tag{3.4}$$

where the two terms are orthogonal if and only if $s\geqq t$; in this case it can be seen that (3.4) is precisely the orthogonal decomposition

$$x(s)=\hat{E}\{x(s)|H_t^-(x)\}+[x(s)-\hat{E}\{x(s)|H_t^-(x)\}]. \tag{3.5}$$

We shall use a symmetric argument to determine the backward representation. More precisely, for $s\leqq t$ we shall derive a backward version of (3.4) from the decomposition

$$x(s)=\hat{E}\{x(s)|H_t^+(x)\}+[x(s)-\hat{E}\{x(s)|H_t^+(x)\}]. \tag{3.6}$$

---

[2] "$H_1\perp H_2$" means "H_1 and H_2 are orthogonal".

In view of the Markov property (4.2) and the standard projection formula [11] the first term in (3.6) can be written

$$\begin{aligned}\hat{E}\{x(s)|H_t^+(x)\} &= E\{x(s)x(t)'\}E\{x(t)x(t)'\}^{-1}x(t)\\ &= Pe^{A'(t-s)}P^{-1}x(t) = e^{-PA'P^{-1}(s-t)}x(t),\end{aligned} \tag{3.7}$$

where we have used (3.4) to evaluate $E\{x(s)x(t)'\}$. From (3.7) it is clear that

$$\xi(t) = e^{PA'P^{-1}t}x(t) \tag{3.8}$$

is a wide sense *backward martingale* with respect to the family $\{H_t^+(x)\}$, i.e.

$$\hat{E}\{\xi(s)|H_t^+(x)\} = \xi(t) \quad \text{for } s \leqq t, \tag{3.9}$$

and using (3.3) we obtain

$$d\xi = e^{PA'P^{-1}t}[(AP+PA')P^{-1}x\,dt + B\,dw],$$

which, because of (1.5), may be written

$$d\xi = e^{PA'P^{-1}t}B(dw - B'P^{-1}x\,dt). \tag{3.10}$$

LEMMA 3.1. *Let $\{x(t);\ t\in R\}$ be the solution on $(-\infty,\infty)$ of (3.3), and let P be the covariance matrix of x. Then the vector process $\bar{w}$, defined by*

$$d\bar{w} = dw - B'P^{-1}x\,dt; \qquad \bar{w}(0)=0, \tag{3.11}$$

belongs to class $\mathscr{W}$, and $H_t^-(d\bar{w})$ is orthogonal to $H_t^+(x)$ for all $t\in R$.

Proof. Inserting (1.6) and (1.7a) for w and x in (3.11) yields

$$\bar{w}(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega} T(i\omega)\,d\hat{w}(\omega). \tag{3.12}$$

where

$$T(s) = I - B'P^{-1}(sI-A)^{-1}B. \tag{3.13}$$

Consequently $\bar{w}$ is a zero-mean, mean-square continuous vector process with stationary increments and spectral density $T(s)T(-s)'$ and such that $\bar{w}(0)=0$. Then, to see that $\bar{w}$ is of class $\mathscr{W}$, it just remains to show that

$$T(s)T(-s)' = I. \tag{3.14}$$

To this end first note that

$$T(s)T(-s)' = I - B'P^{-1}(sI-A)^{-1}B - N(s)P^{-1}B, \tag{3.15}$$

where

$$N(s) = T(s)B'(-sI-A')^{-1} \tag{3.16a}$$

$$= B'(-sI-A')^{-1} - B'P^{-1}(sI-A)^{-1}BB'(-sI-A')^{-1}. \tag{3.16b}$$

In view of (1.5) we may write

$$BB' = (sI-A)P + P(-sI-A')$$

which inserted into (3.16b) yields

$$N(s) = -B'P^{-1}(sI-A)^{-1}P. \tag{3.17}$$

Now (3.15) and (3.17) together yield (3.14). To show that $H_t^-(d\bar{w})\perp H_t^+(x)$, take

$t_1 \leqq t_2 \leqq t_3$ and form

$$E\{[\bar{w}(t_1)-\bar{w}(t_2)]x(t_3)'\} = \int_{-\infty}^{\infty} \frac{e^{i\omega t_1}-e^{i\omega t_2}}{i\omega} e^{-i\omega t_3} N(i\omega)\, d\omega. \tag{3.18}$$

Here we have used (3.12), (1.7a) and (3.16a) to obtain (3.18). But $(e^{-i\omega\alpha}-e^{-i\omega\beta})/i\omega$ is the Fourier transform of the indicator function $\chi_{(\alpha,\beta)}$ of the interval (α, β) and, in view of (3.17), $N(i\omega)$ is the Fourier transform of $-B'P^{-1}e^{At}P\chi_{(0,\infty)}$. Hence Parseval's Theorem yields

$$E\{[\bar{w}(t_1)-\bar{w}(t_2)]x(t_3)'\} = B'P^{-1}\int_{-\infty}^{\infty} \chi_{(t_1-t_3,t_2-t_3)}(t)\chi_{(0,\infty)}(t)\, e^{At}\, dt\, P,$$

which is zero whenever $t_1, t_2 \leqq t_3$. □

Consequently, in view of (3.7)–(3.11), (3.6) can be written

$$\begin{aligned} x(s) &= e^{-PA'P^{-1}(s-t)}x(t)+e^{-PA'P^{-1}s}[\xi(s)-\xi(t)] \\ &= e^{-PA'P^{-1}(s-t)}x(t)+\int_t^s e^{-PA'P^{-1}(s-\tau)}B\, d\bar{w}(\tau), \end{aligned} \tag{3.19}$$

which is the backward counterpart of (3.4). Since $\mathrm{Re}\,\{\lambda(-PA'P^{-1})\}>0$ and $H_t^-(d\bar{w})\perp H_t^+(x)$ for all $t \in R$,

$$dx = -PA'P^{-1}x\, dt + B\, d\bar{w}, \tag{3.20}$$

obtained by differentiating (3.19), is a backward representation of x. In [22], [30] it was shown that, for arbitrary w and $\bar{w}$ of class $\mathscr{W}$, the solutions on $(-\infty, \infty)$ of (3.3) and (3.20) have the same second-order properties. Here we have demonstrated that, for the particular choice (3.11) of $\bar{w}$, these systems actually represent the same wide sense Markov process. We record this observation in the following theorem.

THEOREM 3.2. *Let $\{x(t); t \in R\}$ be a vector-valued, wide sense stationary, purely nondeterministic, wide sense Markov process with covariance matrix P. Then x has a forward representation* (3.3) *with* $\mathrm{Re}\,\{\lambda(A)\}<0$ *and $H_t^+(dw)\perp H_t^-(x)$ for all $t \in R$, and a corresponding backward representation* (3.20) *with $H_t^-(d\bar{w})\perp H_t^+(x)$ for all $t \in R$. The processes x, w and $\bar{w}$ are related as in* (3.11).

In § 1 we only considered stochastic realizations for which $\mathrm{Re}\,\{\lambda(A)\}<0$, i.e. with the state process x written in the forward form. From what has been said above, it is clear that we will get an isomorphic theory by reversing time. In particular, let us consider representations of the type

$$d\bar{x} = \bar{A}\bar{x}\, dt + \bar{B}\, d\bar{w}, \tag{3.21a}$$

$$d\bar{z} = \bar{C}\bar{x}\, dt + \bar{D}\, d\bar{w}, \tag{3.21b}$$

where $\mathrm{Re}\,\{\lambda(\bar{A})\}>0$ and $H_t^-(dw)\perp H_t^+(x)$ for all $t \in R$. We shall call (3.21) a proper or a wide sense *backward stochastic realization* of y, depending on whether the solution $\bar{z}$ of (3.21) on $(-\infty, \infty)$ equals y or has the same spectral density as y. Equation (3.21a) has the unique solution

$$\bar{x}(t) = -\int_t^{\infty} e^{\bar{A}(t-\tau)}\bar{B}\, d\bar{w}(\tau) \tag{3.22}$$

on $(-\infty, \infty)$, and by the procedure used in § 1 we obtain

$$\bar{z}(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega t}-1}{i\omega} \bar{W}(i\omega)\, d\hat{\bar{w}}(\omega) \tag{3.23}$$

where

$$\bar{W}(s)=\bar{C}(sI-\bar{A})^{-1}\bar{B}+\bar{D}. \tag{3.24}$$

If (3.21) is a backward stochastic realization of y, we must have

$$\bar{W}(s)\bar{W}(-s)'=\Phi(s), \tag{3.25}$$

i.e. $\bar{W}$ is a *strictly unstable spectral factor* of Φ. Conversely, each such spectral factor $\bar{W}$ is the transfer function of an equivalence class of wide sense backward stochastic realizations; to see this proceed as in § 1. If $\bar{W}$ is minimal, we shall say that the realization (3.21) is *minimal*; only such representations will be considered in the sequel.

Consider the problem of determining all strictly unstable minimal spectral factors (3.24) of Φ. Since $\bar{W}(-s)\bar{W}(s)'=\Phi(s)'$, this problem is equivalent to finding all stable minimal factors $\bar{W}(-s)$ of $\Phi(s)'$. Given the representation (1.14)–(1.15), we have

$$\Phi(s)'=\bar{Z}(s)+\bar{Z}(-s)', \tag{3.26}$$

where $\bar{Z}$ is the positive real matrix function Z', i.e.

$$\bar{Z}(s)=G'(sI-F')^{-1}H'+\tfrac{1}{2}R. \tag{3.27}$$

Consequently we have reduced the problem to the one considered in § 1. In fact, all stable factors

$$\bar{W}(-s)=\bar{C}(sI+\bar{A})^{-1}(-\bar{B})+\bar{D} \tag{3.28}$$

of $\Phi(s)'$ are given by

$$[-\bar{A},-\bar{B},\bar{C},\bar{D}]=[TF'T^{-1},\ T(-\bar{B}_1,-\bar{B}_2)S,\ G'T^{-1},\ (R^{1/2},0)S] \tag{3.29}$$

where T is any nonsingular $n\times n$-matrix, S is any orthogonal matrix of appropriate dimension and $(\bar{B}_1,\bar{B}_2)$ satisfy

$$F'\bar{P}+\bar{P}F+\bar{B}_1\bar{B}_1'+\bar{B}_2\bar{B}_2'=0, \tag{3.30a}$$

$$\bar{P}G-\bar{B}_1R^{1/2}=H', \tag{3.30b}$$

(3.30c) $\bar{P}$ is a symmetric, positive definite $n\times n$-matrix.

This the *dual spectral factorization problem* considered by Anderson [6] and Faurre [11]. As in the forward setting it is no restriction to take $T=I$ and $S=I$, i.e. to consider backward stochastic realizations of the form $[-F',(\bar{B}_1,\bar{B}_2),G',(R^{1/2},0)]$ only; then $\bar{P}$ in (3.30) is the state covariance matrix.

Let $\bar{\Lambda}$ be given by (2.8) and define $\bar{\mathcal{P}}=\{P=P'|\bar{\Lambda}(P)\leq 0\}$ and $\bar{\mathcal{P}}_0=\{P\in\bar{\mathcal{P}}|\bar{\Lambda}(P)=0\}$. By Theorem 2.1, the set $\bar{\mathcal{P}}$ is closed, bounded and convex, and there are two elements $\bar{P}_*$ and $\bar{P}^*$ in $\bar{\mathcal{P}}_0$ such that $\bar{P}_*\leq P\leq\bar{P}^*$ for all $P\in\bar{\mathcal{P}}$. Moreover, $\bar{\mathcal{P}}$ is the set of all solutions $\bar{P}$ of (3.30), and $\bar{\mathcal{P}}_0$ is the set of all such solutions for which $\bar{B}_2=0$. Let $\bar{\mathcal{B}}$ be the set of all solutions $\bar{B}=(\bar{B}_1,\bar{B}_2)$ of (3.30a)–(3.30b) as $\bar{P}$ varies over $\bar{\mathcal{P}}$, and let $\bar{B}_*$ and $\bar{B}^*$ be the elements in $\bar{B}$ corresponding to $\bar{P}_*$ and $\bar{P}^*$ respectively. As expressed by the following lemma (which is essentially the same as one found in [11]) there is a one-one correspondence between $\mathcal{P}$ and $\bar{\mathcal{P}}$ as well as between $\mathcal{B}$ and $\bar{\mathcal{B}}$.

LEMMA 3.3. *The set of matrices* $(\bar{P},\bar{B}_1,\bar{B}_2)$ *given by*

$$\bar{P}=P^{-1}, \tag{3.31a}$$

$$(\bar{B}_1,\bar{B}_2)=P^{-1}(B_1,B_2) \tag{3.31b}$$

is a solution of (3.30) *if and only if* (P, B_1, B_2) *is a solution of* (1.17). *In particular,* $\bar{P}_* = (P^*)^{-1}$, $\bar{P}^* = (P_*)^{-1}$, $\bar{B}_* = (P^*)^{-1}B^*$ *and* $\bar{B}^* = (P_*)^{-1}B_*$.

Proof. Pre- and postmultiplying (1.17a) by P^{-1} and premultiplying (1.17b) by P^{-1}, it is seen that P is a solution of (1.17) if and only if (3.31a) is a solution of (3.30) with $(\bar{B}_1, \bar{B}_2)$ given by (3.31b). The rest of the statement then follows trivially from (3.31). □

Lemma 3.3 defines a bijective mapping between the sets $\mathcal{B}$ and $\bar{\mathcal{B}}$. This raises the question whether to each *proper* minimal stochastic realization with transfer function W there is a *unique* proper backward minimal stochastic realization whose transfer function is the dual spectral factor $\bar{W}$, and vice versa. In general this is not true, for a spectral factor may correspond to many proper minimal stochastic realizations (Theorem 7.1). However, we shall see that if, in addition, we require that the two realizations have the same *state space*, i.e. $H_t(\bar{x}) = H_t(x)$, for all $t \in R$, there is such a one-one correspondence under mild conditions on B, and that the input processes are related as in Lemma 3.1. Of course, taking (3.31) *and* (3.11) as the starting point, the families of forward and backward proper minimal stochastic realizations are seen to be bijectively related regardless of any condition on B.

THEOREM 3.4. *Let* (F, G, H, R) *be defined as in* § 1. *To each proper minimal stochastic realization of* y *of the form*

$$dx = Fx\,dt + B_1\,du + B_2\,dv, \tag{3.32a}$$

$$dy = Hx\,dt + R^{1/2}\,du, \tag{3.32b}$$

with state covariance matrix P, *there is one and, if* B_2 *has linearly independent columns, only one proper backward minimal stochastic realization of the form*

$$d\bar{x} = -F'\bar{x}\,dt + \bar{B}_1\,d\bar{u} + \bar{B}_2\,d\bar{v}, \tag{3.33a}$$

$$dy = G'\bar{x}\,dt + R^{1/2}\,d\bar{u}, \tag{3.33b}$$

with state covariance $\bar{P}$, *such that* (3.31) *holds and* $H_t(\bar{x}) = H_t(x)$ *for all* $t \in R$. *Conversely, to each realization* (3.33) *there is one and, if* $\bar{B}_2$ *has linearly independent columns, only one realization* (3.32) *such that* (3.31) *holds and* $H_t(x) = H_t(\bar{x})$ *for all* $t \in R$. *The stochastic processes in the two realizations are related in the following way*

$$\bar{x}(t) = P^{-1}x(t), \tag{3.34}$$

$$d\bar{u} = du - B_1'P^{-1}x\,dt; \qquad \bar{u}(0) = 0, \tag{3.35a}$$

$$d\bar{v} = dv - B_2'P^{-1}x\,dt; \qquad \bar{v}(0) = 0. \tag{3.35b}$$

The relations (3.31), (3.34) *and* (3.35) *define a bijective mapping between the families* (3.32) *and* (3.33) *of forward and backward stochastic realizations.*

Proof. The backward representation (3.20) corresponding to (3.32a) is

$$dx = -PF'P^{-1}x\,dt + B_1\,d\bar{u} + B_2\,d\bar{v}, \tag{3.36a}$$

where, according to Theorem 3.2, $\bar{u}$ and $\bar{v}$ are given by (3.35). Then (3.32b) and (3.35a) together yield

$$dy = (HP + R^{1/2}B_1')P^{-1}x\,dt + R^{1/2}\,d\bar{u},$$

which, in view of (1.17b), is the same as

$$dy = G'P^{-1}x\,dt + R^{1/2}\,d\bar{u}. \tag{3.36b}$$

Now let $\bar{x}$ be defined by (3.34). Then $H_t(\bar{x}) = H_t(x)$ for all $t \in R$ and $\bar{x}$ has the covariance matrix (3.31a). Moreover, (3.34) applied to (3.36) yields (3.33) with $\bar{B}$ given by (3.31b). Secondly, consider an arbitrary proper backward minimal realization

$$d\tilde{x} = -F'\tilde{x}\,dt + \bar{B}_1\,d\tilde{u} + \bar{B}_2\,d\tilde{v}, \tag{3.37a}$$

$$dy = G'\tilde{x}\,dt + R^{1/2}\,d\tilde{u} \tag{3.37b}$$

with $\bar{B}$ given by (3.31b) and $H_t(\tilde{x}) = H_t(x)$ for all $t \in R$. Due to the last condition, there is a nonsingular matrix S such that $x(t) = S\tilde{x}(t)$; since x and $\tilde{x}$ are stationary, S is constant. Set $T = P^{-1}S$. Then in view of (3.34), $\bar{x}(t) = T\tilde{x}(t)$. Hence (3.37) can be written

$$d\bar{x} = -TF'T^{-1}\bar{x}\,dt + T\bar{B}_1\,d\tilde{u} + T\bar{B}_2\,d\tilde{v}, \tag{3.38a}$$

$$dy = G'T^{-1}\bar{x}\,dt + R^{1/2}\,d\tilde{u}. \tag{3.38b}$$

Since $\bar{x}$ and $\tilde{x}$ have the same covariance matrix $\bar{P}$, we must have $T\bar{P}T' = \bar{P}$. Hence, in view of (3.38), (3.30) holds also with $(\bar{P}, F', \bar{B}, G')$ exchanged for $(T\bar{P}T', TF'T^{-1}, T\bar{B}, G'T^{-1})$; in particular, (3.30b) yields $T(\bar{P}G' - \bar{B}_1R^{1/2}) = H'$, which together with the original (3.30b) gives us $TH' = H'$. We also have $TF'T^{-1} = F'$. To see this, form $\hat{E}\{\bar{x}(s)|H_t^+(\bar{x})\}$ for all $s \leqq t$ by using first (3.33) and then (3.38); we get $e^{-F'(s-t)}\bar{x}(t)$ and $e^{-TF'T^{-1}(s-t)}\bar{x}(t)$ respectively. Hence $(F')^iH' = T(F')^iT^{-1}H' = T(F')^iH'$ for $i = 1, 2, \cdots, n$, and since (H, F) is observable we must have $T = I$. Therefore $\tilde{x} = \bar{x}$. Then comparing (3.33b) and (3.37b), we see that $\tilde{u} = \bar{u}$, and hence (3.33a) and (3.37a) yield $\tilde{v} = \bar{v}$, for the columns of $\bar{B}$ are linearly independent. Hence (3.33) and (3.37) are identical. Finally, the converse statement is obtained in the same way starting out with the backward realization (3.33). □

4. The minimum- and maximum-variance realizations. The proper stochastic realizations corresponding to P_* and P^*, the minimum and maximum elements of the set $\mathcal{P}$, will play an important role in what follows. Therefore we shall begin by providing an interpretation of these.

Consider an arbitrary proper minimal stochastic realization of the form (3.32) and with state covariance P. It is not hard to see that such a realization exists; we postpone the proof of this to § 7 (Theorem 7.1). It is well-known [35] that, for each fixed $T \in R$, the estimate

$$\hat{x}(t; T) = \hat{E}\{x(t)|H_{[T,t]}(dy)\} \qquad (t \geqq T) \tag{4.1}$$

is generated by the *Kalman–Bucy filter*

$$d\hat{x} = F\hat{x}\,dt + K(t-T)\,d\nu_T; \qquad \hat{x}(T; T) = 0 \qquad (T \leqq t < \infty), \tag{4.2a}$$

where $\{\nu_T(t);\, t \in [T, \infty)\}$ is the transient *innovation process*, defined by[3]

$$d\nu_T = R^{-1/2}(dy - H\hat{x}\,dt); \qquad \nu_T(\max\{0, T\}) = 0. \tag{4.2b}$$

The matrix function K, called the *Kalman–Bucy gain*, can be determined from the matrix Riccati equation

$$\dot{\Sigma} = F\Sigma + \Sigma F' - KK' + BB'; \qquad \Sigma(0) = P, \tag{4.3a}$$

$$K = \Sigma H'R^{-1/2} + B_1. \tag{4.3b}$$

[3] Our choice of initial conditions in (4.2b) and (4.5b), which are otherwise arbitrary, is to insure that $\nu_T(0) = 0$ ($\bar{\nu}_T(0) = 0$) for negative (positive) T.

In the same manner, given an arbitrary proper backward minimal stochastic realization of the form (3.33), it can be seen that

$$\hat{x}_b(t, T) = \hat{E}\{\bar{x}(t)|H_{[t,T]}(dy)\} \qquad (t \leqq T) \tag{4.4}$$

is given by the *backward Kalman–Bucy filter*

$$d\hat{x}_b = -F'\hat{x}_b\,dt + \bar{K}(T-t)\,d\bar{\nu}_T; \qquad \hat{x}_b(T, T) = 0 \qquad (-\infty < t \leqq T), \tag{4.5a}$$

where $\{\bar{\nu}_T(t);\, t \in (-\infty, T]\}$, defined by

$$d\bar{\nu}_T = R^{-1/2}(dy - G'\hat{x}_b\,dt); \qquad \nu_T(\min\{0, T\}) = 0, \tag{4.5b}$$

is the transient *backward innovation process*, introduced in [17]. Here $\bar{K}$ is given by the dual matrix Riccati equation

$$\dot{\bar{\Sigma}} = F'\bar{\Sigma} + \bar{\Sigma}F - \bar{K}\bar{K}' + \bar{B}\bar{B}'; \qquad \bar{\Sigma}(0) = \bar{P}, \tag{4.6a}$$

$$\bar{K} = \bar{\Sigma}GR^{-1/2} - \bar{B}_1. \tag{4.6b}$$

Note that both ν_T and $\bar{\nu}_T$ are normalized orthogonal increment processes [17], so (4.2) and (4.5) can be regarded as a pair of "nonstationary stochastic realizations" of y. We shall now demonstrate that the steady-state versions of these representations are indeed proper stochastic realizations in the sense of this paper.

THEOREM 4.1. *There is one and only one proper stochastic realization* (3.32) *with state covariance matrix* P_*, *namely*

$$dx_* = Fx_*\,dt + B_*\,du_*, \qquad dy = Hx_*\,dt + R^{1/2}\,du_*, \tag{4.7}$$

and it is the steady-state Kalman–Bucy filter in the sense that, for each $t \in R$, $x_*(t)$, $u_*(t)$ *and* B_* *are the limits in mean square of* $\hat{x}(t, T)$, $\nu_T(t)$ *and* $K(t-T)$ *respectively as* $T \to -\infty$. *The* innovation process u_* *satisfies*

$$H_t^-(du_*) = H_t^-(dy) \tag{4.8}$$

for all $t \in R$, *and the projection of the state* $x(t)$ *of any stochastic realization* (3.32) *onto* $H_t^-(dy)$, *being given by*

$$\hat{E}\{x(t)|H_t^-(dy)\} = x_*(t), \tag{4.9}$$

is invariant with respect to the particular realization.

THEOREM 4.2. *There is one and only one proper stochastic realization* (3.32) *with state covariance* P^*, *namely*

$$dx^* = Fx^*\,dt + B^*\,du^*, \qquad dy = Hx^*\,dt + R^{1/2}\,du^*, \tag{4.10}$$

and it is the forward counterpart (in the sense of Theorem 3.4*) of the backward stochastic realization*

$$d\bar{x}_* = -F'\bar{x}_*\,dt + \bar{B}_*\,d\bar{u}_*, \qquad dy = G'\bar{x}_*\,dt + R^{1/2}\,d\bar{u}_* \tag{4.11}$$

where $\bar{x}_*(t)$, $\bar{u}_*(t)$ *and* $\bar{B}_*$ *are the limits in mean square of* $\hat{x}_b(t;T)$, $\bar{\nu}_T(t)$ *and* $\bar{K}(T-t)$ *respectively as* $T \to \infty$. *Then* x^* *and* u^* *are given by*

$$x^*(t) = P^*\bar{x}_*(t), \tag{4.12}$$

$$du^* = d\bar{u}_* - \bar{B}_*'P^*\bar{x}_*\,dt; \qquad u^*(0) = 0 \tag{4.13}$$

and B_* *by Lemma* 3.3. *The* backward innovation process $\bar{u}_*$ *has the property*

$$H_t^+(d\bar{u}_*) = H_t^+(dy) \tag{4.14}$$

for all $t \in R$, *and*

$$\hat{E}\{\bar{x}(t)|H_t^+(dy)\} = \bar{x}_*(t) \tag{4.15}$$

for the state process $\bar{x}$ *of any backward stochastic realization* (3.33).

Before proving these theorems a few remarks are in order:

(i) It is well-known that

$$E\{[x(t)-\hat{x}(t;\,T)][x(t)-\hat{x}(t;\,T)]'\} = \Sigma(t-T), \tag{4.16}$$

where Σ is given by (4.3); the stationarity of x insures that (4.16) depends on the difference $t-T$ only. Likewise, set $E\{\hat{x}(t;\,T)\hat{x}(t;\,T)'\} = \Pi(t-T)$. Then

$$\Sigma(t) = P - \Pi(t). \tag{4.17}$$

Inserting (4.17) into (4.3) and applying (1.17) it is seen that Π satisfies (2.6) and that

$$K = (G - \Pi H')R^{-1/2}. \tag{4.18}$$

Hence $K(t) \to B_*$ as $t \to \infty$ by Theorem 2.2. The corresponding dual results are analogous. Consequently one could base the proofs of Theorems 4.1 and 4.2 on Theorem 2.2, but instead we shall offer a self-contained proof which is more direct. Note that (4.18) together with (2.6), and its dual counterparts, imply that the filters (4.2) and (4.5) are in fact invariant with respect to the particular realization which provides the process $x(\bar{x})$.

(ii) The choice of (3.33) as the standard form for the backward stochastic realizations rather than (3.36) is motivated by the dual spectral factorization problem. Relation (4.15) provides an additional justification for this choice. As in (4.9), the left member of (4.15) is invariant with respect to variations in the state process $\bar{x}$. On the other hand, were we to project the state process x of (3.36) onto the future space $H_t^+(dy)$, we would have

$$\hat{E}\{x(t)|H_t^+(dy)\} = P(P^*)^{-1}x^*(t), \tag{4.19}$$

which does not enjoy the same invariance properties. Indeed the natural setting for the process x is the forward, and not the backward, realization problem.

Proof of Theorem 4.1. For each fixed $t \in R$ the process $\{\xi(\tau);\, \tau \geqq -t\}$, where $\xi(\tau) = \hat{x}(t;\,-\tau)$, is a uniformly integrable wide sense martingale [10], and therefore $\hat{x}(t;\,T)$ tends to a limit $x_*(t)$ in mean square as $T \to -\infty$. Moreover,

$$\hat{x}(t,\,T) = \hat{E}\{x(t)|H_{[T,t]}(dy)\} \to \hat{E}\{x(t)|\textstyle\bigvee_{T \leqq t} H_{[T,t]}(dy)\} \tag{4.20}$$

in mean square [10], and hence (4.9) holds (a.s. for each t), for $\bigvee_{T\leqq t} H_{[T,t]}(dy) = H_t^-(dy)$. Then ν_T tends to a limit process u_*. Since ν_T has normalized orthogonal increments, the same must hold for u_*; hence u_* is of class $\mathcal{W}$. In view of (4.20), $\Pi(t)$ and $K(t)$, as given by (4.17) and (4.18), tend to limits; let us call these Π_∞ and K_∞ respectively. Consequently, x_* and u_* must satisfy

$$dx_* = Fx_*\,dt + K_\infty\,du_*, \qquad dy = Hx_*\,dt + R^{1/2}\,du_*,$$

which is a proper minimal stochastic realization of y with state covariance Π_∞. Thus $\Pi_\infty \in \mathcal{P}$. But since (4.16) is nonnegative definite for all $t \in R$, (4.17) implies that $P \geqq \Pi_\infty$, and this holds for all $P \in \mathcal{P}$, for the realization (3.32) is arbitrary. (By Theorem 7.1 there is a proper stochastic realization for each $P \in \mathcal{P}$.) Therefore $\Pi_\infty = P_*$, and consequently $K_\infty = B_*$. Given P_*, the matrix B_* is uniquely determined by (2.9a). Moreover, as we shall see in § 5, u_* is uniquely determined as a *causal* function of y through relations (5.10b) and (5.12). Hence there is only one proper stochastic realization (3.32) with $P = P_*$, and moreover $H_t^-(du_*) \subset H_t^-(dy)$. Since, in addition, $H_t^-(du_*) \supset H_t^-(dy)$, (4.8) holds. Also, since x_* is uniquely determined, the limit (4.20) is independent of the choice of state process x. □

Proof of Theorem 4.2. The statements concerning (4.11), (4.14) and (4.15) follow along the same lines as in the proof of Theorem 4.1, just reversing time. Then the statements concerning (4.10), (4.12) and (4.13) are a consequence of Theorem 3.4. □

5. Internal stochastic realizations. Consider an arbitrary proper stochastic realization (3.32) and its backward counterpart (3.33). The following lemma describes the relationship between the two input processes w and $\bar{w}$ and the output process y.

LEMMA 5.1. *Let $(w, \bar{w})$ be the pair of input processes defined above. Then the following relations hold for all $t \in R$.*

(i) $H_t^-(dy) \subset H_t^-(dw)$ *and* $H_t^+(dy) \subset H_t^+(d\bar{w})$,

(ii) $H(y) \subset H(w)$,

(iii) $H_t^-(d\bar{w}) \subset H_t^-(dw)$ *and* $H_t^+(dw) \subset H_t^+(d\bar{w})$,

(iv) $H(\bar{w}) = H(w)$.

Proof. Relations (i) and (ii) are trivial consequences of (1.1b) and (1.3) and (3.21b) and (3.22), recalling that $z = \bar{z} = y$. To obtain (iii), insert first (1.3) and then $\bar{x} = P^{-1}x$, as given by (3.22), into (3.11). Then (iv) is proven by letting $t \to \infty$ in the first of relations (iii) and $t \to -\infty$ in the second. □

Since the input process w is of class $\mathcal{W}$, (i) implies that the future increments of w are orthogonal to the past increments of y, i.e. $H_t^+(dw) \perp H_t^-(dy)$ for all $t \in R$. In the same manner it can be seen that $H_t^-(d\bar{w}) \perp H_t^+(dy)$ for all t. It follows from Theorem 5.5 below that the innovation process u_* and the backward innovation process $\bar{u}_*$ are the only input processes to satisfy relations (i) with equality; they satisfy (4.8) and (4.14) respectively. The only thing we can say about the future space of u^* is that $H_t^+(du^*) \subset H_t^+(dy)$, which follows from Theorem 5.5. Hence we have again detected a certain lack of symmetry between the minimum- and maximum-variance realizations.

We shall now consider those realizations for which the converse of relation (ii) holds.

DEFINITIONS. The proper forward or backward stochastic realization $[A, B, C, D; w]$ of y is said to be *internal* if $H(w) = H(y)$. If $H(w) \neq H(y)$, the realization is said to be *external.*

For an internal stochastic realization, the input process w can be expressed in terms of the output y. Therefore, if x is the state process, $x(t) \in H(y)$ for all $t \in (-\infty, \infty)$. In view of Lemma 5.1 (iv), the backward counterpart of any internal (forward) realization is also internal. Hence, in the sequel, we shall restrict our attention to forward realizations, and only consider backward ones when there is an interplay between the forward and backward settings. We now turn to the characterization of the set of internal realizations.

THEOREM 5.2. *A proper stochastic realization of y is internal if and only if it has a square transfer function W, i.e. $W(s)$ is $m \times m$.*

Proof. The proof consists of two parts. First we show that $H(w) = H(y)$ if and only if W has a left inverse. Secondly we show that W has a left inverse if and only if it is $m \times m$.

(i) Assume that $w(t) \in H(y)$ for all $t \in R$. Then there is a representation

$$w(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega t} - 1}{i\omega} K(i\omega)\, d\hat{y}(\omega) \tag{5.1}$$

satisfying the conditions of Lemma 2.3. Therefore, since the stochastic spectral measure is unique, $d\hat{w} = K(i\omega)\,d\hat{y}$. But

$$d\hat{y} = W(i\omega)\,d\hat{w}, \tag{5.2}$$

for $y = z$ satisfies (1.7b), and consequently

$$d\hat{w} = K(i\omega)W(i\omega)\,d\hat{w}. \tag{5.3}$$

Postmultiply (5.3) by $d\hat{w}\dagger$, take expectation, and note that $E\{d\hat{w}\,d\hat{w}\dagger\} = I\,d\omega$ to see that

$$K(s)W(s) = I \tag{5.4}$$

by analytic continuation. Hence W has a left inverse. Conversely, assume that W has a left inverse K. Then (5.3) holds, and, in view of (5.2), we have (5.1). Hence $w(t) \in H(y)$ for all $t \in R$, and therefore $H(w) = H(y)$ (Lemma 5.1 (ii)).

(ii) An $m \times k$ rational transfer matrix $W(s)$ has a left inverse if and only if $\rho\{W\} = k$, where ρ stands for rank, defined with respect to the field of rational functions [34; p. 162, Thm. 5.5.3]. Therefore it remains to show that $\rho\{W\} = k$ if and only if $k = m$. To this end, apply Sylvester's inequality [34; p. 40] to (1.10) to obtain

$$\rho\{W(s)\} + \rho\{W(-s)'\} - k \leqq \rho\{\Phi\} \leqq \min\,[\rho\{W(s)\}, \rho\{W(-s)'\}],$$

which can be written

$$2\rho\{W\} - k \leqq m \leqq \rho\{W\}, \tag{5.5}$$

for $\rho\{\Phi\} = m$. Consequently, if $\rho\{W\} = k$, we have $k = m$. Conversely, if $k = m$, (5.5) implies that $\rho\{W\} = k$. □

COROLLARY 5.3. *A proper minimal stochastic realization in the standard form* (3.32) *is internal if and only if* $B_2 = 0$.

Proof. The transfer function of (3.32) is

$$W(s) = H(sI - F)^{-1}(B_1, B_2) + (R^{1/2}, 0), \tag{5.6}$$

which is square if and only if $B_2 = 0$. □

Consequently the internal stochastic realizations in standard form are precisely the representations of the type

$$dx = Fx\,dt + B\,du, \tag{5.7a}$$

$$dy = Hx\,dt + R^{1/2}\,du \tag{5.7b}$$

among which we have the minimum-variance realization (4.7) and the maximum-variance realization (4.10).

THEOREM 5.4. *There is a one-one correspondence between the family of internal realizations* (5.7) *and the set* $\mathcal{P}_0$ *of solutions of the algebraic Riccati equation* $\Lambda(P) = 0$. *The input process u of* (5.7) *is given by*

$$u(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega t} - 1}{i\omega} W^{-1}(i\omega)\,d\hat{y}, \tag{5.8}$$

where W is the transfer function of (5.7).

Proof. Each stochastic realization (5.7) has a state covariance matrix P which belongs to $\mathcal{P}_0$, since $B_2 = 0$. (Theorem 2.1). Hence it remains to show that to each $P \in \mathcal{P}_0$ there is one and only one proper stochastic realization (5.7) and that u is given by (5.8): To each $P \in \mathcal{P}_0$ there is one and only one spectral factor of the form (5.6), namely the square factor

$$W(s) = H(sI - F)^{-1}B + R^{1/2}, \tag{5.9}$$

for B is uniquely determined by (2.9a). Since R is nonsingular, (5.9) has an inverse W^{-1}. First define u by (5.8). Then $d\hat{y} = W(i\omega)\, d\hat{u}$, which transformed to the time domain yields (5.7). Secondly, let u be the input process of a proper stochastic realization with transfer function (5.9). Then $d\hat{y} = W(i\omega)\, d\hat{u}$, and hence u is given by (5.8). □

The internal realization (5.7) can be inverted in the time domain also by rewriting it in the form

$$dx = \Gamma x\, dt + BR^{-1/2}\, dy, \tag{5.10a}$$

$$du = R^{-1/2}(dy - Hx\, dt) \tag{5.10b}$$

where, in view of (2.9a),

$$\Gamma = F - BR^{-1/2}H \tag{5.11}$$

is the feedback matrix (2.3). Once there is a solution of (5.10a), u is given by (5.10b). For the two extreme realizations, corresponding to P_* and P^*, such solutions are immediate, namely

$$x_*(t) = \int_{-\infty}^{t} e^{\Gamma_*(t-\tau)} B_* R^{-1/2}\, dy(\tau) \tag{5.12}$$

and

$$x^*(t) = -\int_{t}^{\infty} e^{\Gamma^*(t-\tau)} B^* R^{-1/2}\, dy(\tau) \tag{5.13}$$

respectively. In fact, all eigenvalues of Γ_* (Γ^*) have negative (positive) real parts. (See § 2.) Then u_* and u^* can be determined from (5.10b).

Other internal stochastic realizations can now be handled by integrating stable modes over the past and unstable over the future, provided that the matrix Γ has no eigenvalues on the imaginary axis. However, since $P_* < P^*$ [32, p. 260], no such eigenvalues occur for $\mathcal{P}_0$-realizations [33, p. 630; Remark 19]. In fact the solution is surprisingly simple.

THEOREM 5.5. *Consider an internal stochastic realization* (5.7). *Let* $\Pi^+(\Pi^-)$ *be the projection operator onto the invariant subspace spanned by the eigenvectors corresponding to eigenvalues of the feedback matrix* (5.11) *with positive* (*negative*) *real parts. Then*

$$x(t) = \Pi^- x_*(t) + \Pi^+ x^*(t), \tag{5.14}$$

where x_* *and* x^* *are given by* (5.12) *and* (5.13). *The input process* u *is given by*

$$du = R^{-1/2}[dy - H\Pi^- x_*(t)\, dt - H\Pi^+ x^*(t)\, dt]. \tag{5.15}$$

The proof of Theorem 5.5 is based on the following lemma.

LEMMA 5.6 (J. C. Willems). *Let* $P \in \mathcal{P}_0$, *and let* Π^+ *and* Π^- *be defined as in Theorem* 5.5. *Then* $\Pi^+ + \Pi^- = I$ *and*

$$P = \Pi^- P_* + \Pi^+ P^*. \tag{5.16}$$

Moreover, with Γ_* *and* Γ^* *defined as above,*

$$\Pi^-\Gamma_*\Pi^- = \Pi^-\Gamma_* \quad \textit{and} \quad \Pi^+\Gamma^*\Pi^+ = \Pi^+\Gamma^*. \tag{5.17}$$

In view of the fact that $P^* - P_* > 0$ and (H, F) is observable (see § 1), this result is an immediate consequence of Theorem 6 and Lemma 8 in [33].

Proof of Theorem 5.5. Let P be the state covariance matrix of the stochastic realization (5.7). Hence $P \in \mathcal{P}_0$ (Corollary 5.3). Since $(\Pi^-)^2 = \Pi^-$ and $\Pi^-\Pi^+ = 0$, we have $\Pi^- P = \Pi^- P_*$ from (5.16). Consequently, in view of (2.9a) and (5.11),

$$\Pi^- B = \Pi^- B_*, \tag{5.18a}$$

$$\Pi^- \Gamma = \Pi^- \Gamma_* = \Pi^- \Gamma_* \Pi^-, \tag{5.18b}$$

where in the last relation we have also used (5.17). Hence, premultiplying (5.10a) by Π^- and using (5.18), it is seen that $\Pi^- x(t)$ satisfies the differential equation

$$d\xi = \Pi^- \Gamma_* \xi \, dt + \Pi^- B_* R^{-1/2} \, dy \tag{5.19}$$

on $(-\infty, \infty)$. But $\Pi^- x_*(t)$, too, satisfies (5.19) on $(-\infty, \infty)$. To see this, use (5.17). Therefore, since (5.19) has a unique solution on $(-\infty, \infty)$, we must have $\Pi^- x(t) = \Pi^- x_*(t)$ for all $t \in R$. In the same way we show that $\Pi^+ x(t) = \Pi^+ x^*(t)$. Hence, (5.14) follows from $\Pi^+ + \Pi^- = I$ (Lemma 5.6). Then insert (5.14) into (5.10b) to obtain (5.15). □

It follows from (5.12) and (5.13) that $x_*(t) \in H_t^-(dy)$ and $x^*(t) \in H_t^+(dy)$ for each $t \in R$. Therefore, (5.14) decomposes $x(t) \in H(y)$ into two components, one in $H_t^-(dy)$ and one in $H_t^+(dy)$. In view of (4.8) and (4.14), we can acquire symmetry between past and future by using (4.12) to rewrite (5.14) in the form

$$x(t) = \Pi^- x_*(t) + \Pi^+ P^* \bar{x}_*(t). \tag{5.20}$$

Consequently, the state process of any internal stochastic realization can be expressed in terms of the steady-state forward and backward Kalman–Bucy estimates, x_* and $\bar{x}_*$, and therefore it can be constructed from a linear combination of the filters (4.2) and (4.5), by taking the limit in quadratic mean.

6. Families of totally ordered stochastic realizations. Considering minimal stochastic realizations in the standard form (3.32) leaves only the matrix $B = (B_1, B_2)$ and the input process $w = \binom{u}{v}$ to be determined, the parameters (F, G, H, R) being given. This section will be devoted to studying the set $\mathcal{B}$ of feasible matrices B, defined in § 2; finding w will be the topic of § 7.

It was shown in § 4 (Theorem 4.1) that

$$B_* = \lim_{t \to \infty} K(t), \tag{6.1}$$

where K is the Kalman–Bucy gain function. This fact together with the following theorem provide us with a means to determine B_* directly without first having to obtain P_*.

THEOREM 6.1 (Kailath–Lindquist). *Let (K, Q) be the unique solution on $[0, \infty)$ of the system of matrix differential equations*

$$\dot{K} = -QQ'H'R^{-1/2}; \qquad K(0) = GR^{-1/2} \tag{6.2a}$$

$$\dot{Q} = (F - KR^{-1/2}H)Q; \qquad Q(0) = GR^{-1/2}. \tag{6.2b}$$

Then K is the Kalman–Bucy gain function. The filter covariance function Π, defined in § 4 (Remark (i)), satisfies

$$\dot{\Pi} = QQ'; \qquad \Pi(0) = 0. \tag{6.3}$$

Note that, although different realizations (3.32) yield different Riccati equations (4.3) [but the same filter (4.2)], the *non-Riccati algorithm* (6.2) is invariant over $\mathcal{P}$, depending only on the known quantities (F, G, H, R). If needed, P_* can be determined as the limit of $\Pi(t)$ as $t \to \infty$ (Theorem 2.2), where Π is generated by either (2.6) or (6.3). The system (6.2)–(6.3) is precisely the algorithm derived in [17] by using the transient backward innovation process (4.5b) and in [16] by factoring the matrix differential equation (4.3). A dual non-Riccati algorithm generating the backward Kalman–Bucy gain $\bar{K}$ and the backward filter covariance $\bar{\Pi}$ can be derived analogously by using the *forward* innovation (4.2b) or alternatively from (4.7) by applying the technique of [16]; formally it can be obtained by merely exchanging (F, G, H, R) for (F', G', H', R) in (6.2).

It can be seen that $K(t)$ approaches B_* from outside of $\mathscr{B}$. In fact, as one can see by comparing (2.9a) and (4.18), $K(t)$ is related to $\Pi(t)$ as B_* to P_*, and, in view of (6.3), Π is monotonely nondecreasing starting out with $0 \notin \mathscr{P}$ at $t=0$; hence $\Pi(t) \leqq P_*$ for all t. Here we shall show that there are equations similar to (6.2) whose trajectories, with the proper initial conditions, lie entirely inside $\mathscr{B}$. These equations will consequently generate families of wide sense stochastic realizations. Again the basic idea is to eliminate the need of going via the auxilliary quantity P.

THEOREM 6.2. *Let $[F, B_0, H, (R^{1/2}, 0)]$ be an arbitrary wide sense minimal stochastic realization of y in standard form, and let $\theta \to B(\theta) = [B_1(\theta), B_2(\theta)]$ be the unique solution on $(-\infty, \infty)$ of the system of matrix differential equations*

$$\frac{dB_1}{d\theta} = B_2 B_2' H' R^{-1/2}, \tag{6.4a}$$

$$\frac{dB_2}{d\theta} = (F - B_1 R^{-1/2} H) B_2 \tag{6.4b}$$

with initial condition $B(0) = B_0$. For each $\theta \in (-\infty, \infty)$, let $P(\theta)$ be the unique solution of the Lyapunov equation

$$FP + PF' + B(\theta)B(\theta)' = 0. \tag{6.5}$$

Then, for each $\theta \in (-\infty, \infty)$, $[F, B(\theta), H, (R^{1/2}, 0)]$ is a wide sense minimal stochastic realization of y with state covariance matrix $P(\theta)$. This family of realizations is totally ordered in the sense that $P(\theta_2) \leqq P(\theta_1)$ for $\theta_1 \leqq \theta_2$. If $B_0 \in \mathscr{B}_-$, $B(\theta) \to (B_*, 0)$ as $\theta \to \infty$, and if $B_0 \in \mathscr{B}_+$, $B(\theta) \to (B^*, 0)$ as $\theta \to -\infty$. The function $\theta \to P(\theta)$ satisfies the differential equations (6.7) *and*

$$\frac{dP}{d\theta} = -B_2 B_2', \tag{6.6}$$

and also conditions (iii) *and* (iv) *of Lemma* 6.3 *where here P_0 may be any point on the trajectory $\{P(\theta); -\infty < \theta < \infty\}$.*

The proof of this theorem is based on the following lemma.

LEMMA 6.3. *Let Λ be defined by* (2.1). *Then, for each $P_0 \in \mathscr{P}$, the matrix differential equation*

$$\frac{dP}{d\theta} = \Lambda(P(\theta)); \qquad P(0) = P_0 \tag{6.7}$$

has a unique solution on $(-\infty, \infty)$, such that (i) *$P(\theta) \in \mathscr{P}$ for all $\theta \in (-\infty, \infty)$,* (ii) *$P(\theta_2) \leqq P(\theta_1)$ for $\theta_1 \leqq \theta_2$,* (iii) *if $P_0 \in \mathscr{P}_-$, $P(\theta) \to P_*$ as $\theta \to \infty$, and* (iv) *if $P_0 \in \mathscr{P}_+$, $P(\theta) \to P^*$ as $\theta \to -\infty$.*

Proof. First note that (6.7) can be replaced by the system

$$\frac{dP}{d\theta} = U(\theta)\Lambda(P_0)U(\theta)'; \qquad P(0) = P_0, \tag{6.8a}$$

$$\frac{dU}{d\theta} = \Gamma(\theta)U(\theta); \qquad U(0) = I, \tag{6.8b}$$

where $\Gamma(\theta)$ is the feedback matrix (2.3) corresponding to $P(\theta)$. To see this, reformulate (6.7) to read

$$\frac{dP}{d\theta}=(F-GR^{-1}H)P+P(F-GR^{-1}H)'+PH'R^{-1}HP+GR^{-1}G,$$

and use the differentiation technique employed by Kailath in [15], i.e. observe that

$$\frac{d^2P}{d\theta^2}=\Gamma(\theta)\frac{dP}{d\theta}+\frac{dP}{d\theta}\Gamma(\theta)'; \qquad \frac{dP}{d\theta}(0)=\Lambda(P_0),$$

and integrate to obtain (6.8).

Clearly the Riccati equation (6.7) has a unique solution locally in the neighborhood of $\theta=0$. In fact, at least for small θ, $P(\theta)=Y(\theta)X(\theta)^{-1}$, where the $n\times n$-matrix valued functions X and Y satisfy a system of linear differential equations such that $X(\theta)^{-1}$ exists for sufficiently small θ [8, p. 156]. Since $P_0\in\mathcal{P}$, $\Lambda(P_0)\leqq 0$, and hence, in view of (6.8a), the condition

$$\text{(6.9)} \qquad \frac{dP}{d\theta}\leqq 0$$

holds along this trajectory. Consequently, (6.7) implies $\Lambda(P(\theta))\leqq 0$, i.e. the trajectory is contained in the bounded (Theorem 2.1) set $\mathcal{P}$. Hence the solution can be extended to the whole real line, for $P(\theta)$ will never leave $\mathcal{P}$. Since Λ is locally Lipschitz, this solution is unique. This also proves (i), and (ii) is a consequence of (6.9).

To prove (iv) we use an argument similar to that in Willems [33, p. 631]. In view of the fact that $\Lambda(P_*)\leqq 0$, $S(\theta):=P(\theta)-P_*$ is the solution of

$$\frac{dS}{d\theta}=\Gamma_*S+S\Gamma_*'+SH'R^{-1}HS; \qquad S(0)=P_0-P_*.$$

Since $S(0)>0$ (for $P_0\in\mathcal{P}_+$) and $dS/d\theta\leqq 0$ (by (6.9)), $S(\theta)>0$ for $\theta\leqq 0$. Consequently S^{-1} exists on $(-\infty, 0]$. Let M_* be defined as in Theorem 2.2, and define $V:=S^{-1}-M_*(0)$. It is easy to see that V satisfies

$$\frac{dV}{d\theta}=-\Gamma_*'V-V\Gamma_*$$

on $(-\infty, 0]$. Since $\operatorname{Re}\{\lambda(-\Gamma_*)\}>0$, $V(\theta)\to 0$ as $\theta\to-\infty$, and hence $S(\theta)\to[M_*(0)]^{-1}=P^*-P_*$ (Theorem 2.2). Therefore $P(\theta)\to P^*$ as $\theta\to-\infty$. This proves (iv). The proof of (iii) is analogous; just exchange substar ($_*$) by superstar (*) everywhere and $(-\infty, 0]$ for $[0,\infty)$. (Now $S(0)<0$ for $\theta\geqq 0$.) □

Hence, given any P_0 in $\mathcal{P}_+\cap\mathcal{P}_-$, we may construct a trajectory $\mathcal{T}\subset\mathcal{P}$ extending from P^* through P_0 to P_* so that $\mathcal{T}$ is a totally ordered set of matrices P satisfying (1.17). The only difference between (2.6) and (6.7) is the initial conditions ($0\notin\mathcal{P}$); the differential equation is the same. Its critical points are precisely the elements of $\mathcal{P}_0$, one of which (P_*) is locally stable in the forward direction and another of which (P^*) is stable in the backward direction (cf. [33]). Note, however, that (6.2) and (6.4) are not exactly the same, although they are derived from the same differential equation. A dual (backward) version of (6.1) can be obtained by factoring (2.7), with $\Pi(0)\in\bar{\mathcal{P}}$, as above.

Proof of Theorem 6.2. Let P_0 be the state covariance of the initial realization $[F, B_0, H, (R^{1/2}, 0)]$, and let $\{P(\theta);\ -\infty<\theta<\infty\}$ be the trajectory through P_0 defined by

LEMMA 6.3. Define $B(\theta)$ as

$$B_1(\theta)=[G-P(\theta)H']R^{-1/2}, \tag{6.10a}$$

$$B_2(\theta)=U(\theta)(B_0)_2, \tag{6.10b}$$

where U is given by (6.8b). Then (6.6) and (6.4a) follow from (6.8a) (for $\Lambda(P_0)=-(B_0)_2(B_0)_2'$) and (6.4b) is a consequence of (6.8b) and (6.10). A local Lipschitz condition insures uniqueness. In view of (6.6) and (6.7), we have $B_2(\theta)B_2(\theta)'=-\Lambda(P(\theta))$, which together with (6.10a) yields (6.5). Since Re $\{\lambda(F)\}<0$ and $(F, B(\theta))$ is controllable (for (F, B_0) is), (6.5) has a unique positive definite, symmetric solution [8]. This fact together with (6.5) and (6.10a) insures that $(P(\theta), B(\theta))$ satisfies (1.17), and consequently $[F, B(\theta), H, (R^{1/2}, 0)]$ is a wide sense stochastic realization with state covariance $P(\theta)$. By Lemma 6.2, $P(\theta)$ satisfies conditions (ii)–(iv), and obviously the last two conditions hold for any P_0 on the trajectory $\{P(\theta); -\infty<\theta<\infty\}$. Finally, the fact that $B_1(\theta)$ tends to $B_*(B^*)$ as $\theta\to\infty(\theta\to-\infty)$ under the stated conditions, follows from conditions (iii) and (iv) and (6.10a). Since $dP/d\theta\to 0$, (6.6) implies that $B_2(\theta)\to 0$ as $\theta\to\pm\infty$. □

In the next section we shall interpret Theorem 6.2 in terms of proper stochastic realizations.

7. External stochastic realizations. The following theorem gives a complete characterization of all proper minimal stochastic realizations.

THEOREM 7.1. *Let*

$$dx=Fx\,dt+B_1\,du+B_2\,dv, \tag{7.1a}$$

$$dy=Hx\,dt+R^{1/2}\,du \tag{7.1b}$$

be a proper minimal stochastic realization of y, *and let* $W_1(s)$ *and* $W_2(s)$ *be defined by*

$$W_1(s)=H(sI-F)^{-1}B_1+R^{1/2}, \tag{7.2a}$$

$$W_2(s)=H(sI-F)^{-1}B_2. \tag{7.2b}$$

Then

$$W(s)=[W_1(s), W_2(s)] \tag{7.3}$$

is a minimal stable spectral factor of the spectral density Φ *of* y, *and the input processes are given by*

$$v(t)=\int_{-\infty}^{\infty}\frac{e^{i\omega t}-1}{i\omega}W_2(-i\omega)'\Phi^{-1}(i\omega)\,d\hat{y}(\omega)+z(t) \tag{7.4a}$$

$$u(t)=\int_{-\infty}^{\infty}\frac{e^{i\omega t}-1}{i\omega}W_1(-i\omega)'\Phi^{-1}(i\omega)\,d\hat{y}(\omega) -\int_{-\infty}^{\infty}\frac{e^{i\omega t}-1}{i\omega}W_1^{-1}(i\omega)W_2(i\omega)\,d\hat{z}(\omega) \tag{7.4b}$$

where z *is a mean-square continuous, purely nondeterministic stochastic vector process with stationary increments, zero mean, spectral density*

$$\Psi(s)=I-W_2(-s)'\Phi^{-1}(s)W_2(s), \tag{7.5}$$

and $z(0)=0$. *Moreover,* $\Psi(i\omega)>0$ *for all real* ω *and* $H(z)\perp H(y)$; *we shall call* z *the* exogeneous input component. *Conversely, for each minimal stable spectral factor* (7.3) *of Φ, there is a minimal proper stochastic realization* (7.1) *with u and v given by* (7.4), *z being an arbitrary stochastic vector process with all the properties prescribed above.*

Proof. It was shown in § 1 that, with (7.1) given, (7.3) is a minimal stable spectral factor of Φ; this result is restated here for completeness only. To see that u and v are given by (7.4), first decompose v as

$$v(t)=\hat{E}\{v(t)|H(y)\}+z(t). \tag{7.6}$$

Then $H(z)\perp H(y)$. Given the properties of v and y described in § 1, it is easy to see that the first term in this decomposition is a mean-square continuous, purely nondeterministic vector process with stationary increments, so the same must hold for z; in addition, z has zero mean and $z(0)=0$. Hence, since

$$d\hat{y}(\omega)=W_*(i\omega)\,d\hat{u}_*(\omega), \tag{7.7}$$

where $d\hat{u}_*$ is the stochastic spectral measure of the innovation process u_* and W_* is the transfer function of (4.7), and in view of Lemma 2.3, (7.6) can be written

$$v(t)=\int_{-\infty}^{\infty}\frac{e^{i\omega t}-1}{i\omega}Z(i\omega)\,d\hat{u}_*(\omega)+\int_{-\infty}^{\infty}\frac{e^{i\omega t}-1}{i\omega}\,d\hat{z}(\omega), \tag{7.8}$$

for some Z to be determined. Let Ψ denote the spectral density of the process z. Clearly there is a representation

$$d\hat{z}(\omega)=T(i\omega)\,d\hat{\mu}(\omega), \tag{7.9}$$

where $d\hat{\mu}$ is the stochastic spectral measure of a process μ of classs $\mathcal{W}$ such that $H(\mu)\perp H(y)$, and $T(s)$ is a spectral factor of $\Psi(s)$. Then (7.8) can be written

$$d\hat{v}=Z(i\omega)\,d\hat{u}_*+T(i\omega)\,d\hat{\mu}. \tag{7.10a}$$

Therefore, inserting (7.7) and (7.10a) into

$$d\hat{y}=W_1(i\omega)\,d\hat{u}+W_2(i\omega)\,d\hat{v}, \tag{7.11}$$

which is (7.1) rewritten in terms of spectral measures, and solving for $d\hat{u}$, we obtain

$$d\hat{u}=X(i\omega)\,d\hat{u}_*+Y(i\omega)T(i\omega)\,d\hat{\mu}, \tag{7.10b}$$

where

$$X(s)=W_1^{-1}(s)W_*(s)-W_1^{-1}(s)W_2(s)Z(s) \tag{7.12}$$

and

$$Y(s)=-W_1^{-1}(s)W_2(s), \tag{7.13a}$$

for the matrix R being nonsingular insures that W_1 has an inverse. Since both $\binom{u}{v}$ and $\binom{u_*}{\mu}$ are vector processes of class $\mathcal{W}$, the coefficient matrix function of (7.10), i.e.

$$K(s)=\begin{bmatrix}X(s) & Y(s)T(s)\\ Z(s) & T(s)\end{bmatrix},$$

satisfies relation (2.13) of Lemma 2.3, i.e.

$$X(s)X(-s)'+Y(s)T(s)T(-s)'Y(-s)'=I, \tag{7.14a}$$

$$X(s)Z(-s)'+Y(s)T(s)T(-s)'=0, \tag{7.14b}$$

$$Z(s)Z(-s)'+T(s)T(-s)'=I. \tag{7.14c}$$

Then inserting (7.12) into (7.14b) and applying (7.14c), we have

$$(7.13b)\qquad Z(s)=W_2(-s)'W_*^{-1}(-s)',$$

which inserted into (7.12) yields

$$(7.13c)\qquad X(s)=W_1(-s)'W_*^{-1}(-s)'.$$

To obtain this, we have used the fact that

$$(7.15)\qquad \Phi(s)=W_1(s)W_1(-s)'+W_2(s)W_2(-s)'.$$

Now (7.10) together with (7.7) and (7.13) yield (7.4), and (7.13b) and (7.14c) give us (7.5), for $T(s)T(-s)'=\Psi(s)$. By using the matrix inversion lemma [14, p. 124], we can see that

$$(7.16)\qquad \Psi(s)=[I+W_2(-s)'W_1^{-1}(-s)W_1^{-1}(s)W_2(s)]^{-1}.$$

Hence $\Psi(i\omega)>0$ for all real ω.

Secondly, assume that a minimal stable spectral factor (7.3) is given; from it we can determine a quadruplet $[F, (B_1, B_2), H, (R^{1/2}, 0)]$. Let z be an arbitrary mean-square continuous process with stationary increments, zero mean, and spectral density (7.5), and such that $z(0)=0$ and $H(z)\perp H(y)$. Since z has a rational spectral density, it is purely nondeterministic [9]. Define u and v by (7.4). Then the corresponding stochastic spectral measures $d\hat{u}$ and $d\hat{v}$ are given by (7.10) with X, Y, Z and T defined by (7.13) and (7.9). Straightforward calculations using (7.15) show that X, Y, Z and T satisfy (7.14), and consequently $\binom{u}{v}$ is a process of class $\mathcal{W}$. Finally, with the help of (7.15), we can see that $d\hat{u}$ and $d\hat{v}$ thus defined satisfy (7.11) (the z-components cancel), and therefore (7.1) is a proper stochastic realization of y. □

Theorem 7.1 provides us with an alternative proof of the "only if" part of Corollary 5.3. (Theorem 5.5 gives an alternative proof of the "if" part.) In fact, since $\Psi(i\omega)>0$ for all real ω, the exogeneous input component z is never identically zero. Therefore, unless $B_2=0$, the output of (7.1) contains a component orthogonal to $H(y)$.

We are now in a position to interpret Theorem 6.2 in terms of proper minimal stochastic realizations. Consider an arbitrary such realization

$$(7.17)\qquad dx=Fx\,dt+(B_0)_1\,du_0+(B_0)_2\,dv_0,\qquad dy=Hx\,dt+R^{1/2}\,du_0$$

with exogeneous input component z_0 having spectral density $\Psi_0(s)$. Let $T_0(s)$ be a square spectral factor of $\Psi_0(s)$ and define

$$(7.18)\qquad \mu(t)=\int_{-\infty}^{\infty}\frac{e^{i\omega t}-1}{i\omega}T_0^{-1}(i\omega)\,d\hat{z}_0(\omega).$$

(Since $\Psi_0(i\omega)>0$ for all ω, $T_0(s)$ has an inverse.) Then, $\mu\in\mathcal{W}_k$, where k is the number of columns of $(B_0)_2$. Let $\mathcal{F}$ be the sigma-algebra generated by $\{y(t), \mu(t); t\in R\}$ and form the probability space $(\Omega, \mathcal{F}, P)$ on which (7.17) is defined. Then (7.17) gives rise to a family of proper minimal stochastic realizations

$$(7.19)\qquad dx_\theta=Fx_\theta\,dt+B_1(\theta)\,du_\theta+B_2(\theta)\,dv_\theta,\qquad dy=Hx_\theta\,dt+R^{1/2}\,du_\theta,$$

which are defined on the same probability space $(\Omega, \mathcal{F}, P)$ and which are totally ordered in the sense that the state covariance function $P(\theta)=E\{x_\theta(t)x_\theta(t)'\}$ is monotonely nonincreasing in θ. In fact, for each $\theta\in[-\infty,\infty]$, define $W_1(s;\theta)$ and $W_2(s;\theta)$ by inserting $[B_1(\theta), B_2(\theta)]$, generated by (6.4), into (7.2), and let

$$(7.20)\qquad z_\theta(t)=\int_{-\infty}^{\infty}\frac{e^{i\omega t}-1}{i\omega}T_\theta(i\omega)\,d\hat{\mu}(\omega),$$

where $T_\theta(s)$ is a square spectral factor of

$$\Psi_\theta(s) = I - W_2(-s;\theta)'\Phi^{-1}(s)W_2(s;\theta). \tag{7.21}$$

(We may for example take all T_θ to be minimum phase.) Then define u_θ and v_θ by inserting $W_1(s;\theta)$, $W_2(s;\theta)$ and z_θ into (7.4). Hence $x_\theta(t)$, $u_\theta(t)$ and $v_\theta(t)$ belong to $H(y,\mu)$ for all t and all θ. If $B_0 \in \mathscr{B}_-$, the family (7.19) will contain the steady-state Kalman–Bucy filter (4.7); if $B_0 \in \mathscr{B}_+$, it will contain the maximum-variance model (4.10). Finally, if $B_0 \in \mathscr{B}_0$, (7.19) will only contain one realization, (7.17) itself.

REFERENCES

[1] H. Akaike, *Markovian representation of stochastic processes by canonical variables*, this Journal, 13 (1975), pp. 162–173.
[2] ———, *Stochastic theory of minimal realization*, IEEE Trans. Automatic Control, AC-19 (1974), pp. 667–674.
[3] B. D. O. Anderson, *An algebraic solution to the spectral factorization problem*, Ibid., Automatic Control AC-12 (1967), pp. 410–414.
[4] ———, *A system theory criterion for positive real matrices*, this Journal, 5 (1967), pp. 171–182.
[5] ———, *The inverse problem of stationary covariance generation*, J. Statis. Phys., 1 (1969), pp. 133–147.
[6] ———, *Dual form of a positive real lemma*, Proc. IEEE, 55 (1967), pp. 1749–1750.
[7] B. D. O. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis*, Prentice-Hall, Englewood Cliffs, NJ, 1973.
[8] R. W. Brockett, *Finite Dimensional Linear Systems*, Wiley, New York, 1970.
[9] H. Cramer and M. R. Leadbetter, *Stationary and Related Stochastic Processes*, Wiley, New York, 1967.
[10] J. L. Doob, *Stochastic Processes*, Wiley, New York, 1953.
[11] P. Faurre, *Realisations markoviennes de processus stationnaires*, Research Report no. 13, March 1973, IRIA (LABORIA), Le Chesnay, France.
[12] I. I. Gikhman and A. V. Skorokhod, *Introduction to the Theory of Random Processes*, *W. B.* Saunders, Philadelphia, 1969.
[13] B. L. Ho and R. E. Kalman, *Effective construction of linear state-variable models from input/output functions*, Proc. 3rd Allerton Conf. Circuits and Systems Theory, 1965, pp. 449–459.
[14] A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.
[15] T. Kailath, *Some Chandrasekhar-type algorithms for quadratic regulators*, Proc. IEEE Decision and Control Conference (New Orleans, Dec. 1972).
[16] ———, *Some new algorithms for recursive estimation in constant linear systems*, IEEE Trans. Information Theory IT-19 (1973), pp. 750–760.
[17] A. Lindquist, *Optimal filtering of continuous-time stationary processes by means of the backward innovation process*, this Journal, 12 (1974), pp. 747–754.
[18] ———, *On Fredholm integral equations, Toeplitz equations and Kalman–Bucy filtering*, Appl. Math. Optimization, 1 (1975), pp. 355–373.
[19] ———, *Linear least-squares prediction based on covariance data from stationary processes with finite-dimensional realizations*, Proc. Second European Congress on Operations Research (Stockholm, Sweden), North-Holland, Amsterdam, 1976, pp. 281–286.
[20] A. Lindquist and G. Picci, *A note on the stochastic realization problem*, Proc. Intern. Conf. Information Sciences and Systems (Patras, Greece, Aug. 1976), Hemisphere Publishing Corporation, 1977, pp. 1–5.
[21] ———, *On the structure of minimal splitting subspaces in stochastic realization theory*, Proc. 1977 Decision and Control Conference (New Orleans), pp. 42–48.
[22] L. Ljung and T. Kailath, *Backwards markovian models for second-order stochastic processes*, IEEE Trans. Information Theory, IT-22 (1976), pp. 488–491.
[23] G. Picci, *Stochastic realization of Gaussian processes*, Proc. IEEE, 64 (1976), pp. 112–122.
[24] ———, *Some connections between the theory of sufficient statistics and the identifiability problem*, SIAM J. Appl. Math., 33 (1977), pp. 383–398.
[25] V. M. Popov, *Hyperstability of Control Systems*, Springer-Verlag, New York, 1973.
[26] Yu. A. Rozanov, *On two selected topics connected with stochastic systems theory*, Appl. Math. Optimization, 3 (1976), pp. 73–80.

[27] G. RUCKEBUSCH, *Representations markoviennes de processus gaussiens stationnaires*, Thèse 3ème cycle, Paris VI, 1975.

[28] ———, *Representations markoviennes de processus gaussiens stationnaires et applications statistiques*, Centre de Mathematique Appliquées, internal report no. 18, Ecole Polytechnique, Palaiseau, France.

[29] ———, *On the theory of markovian representation*, preprint.

[30] G. S. SIDHU AND U. A. DESAI, *New smoothing algorithms based on reversed-time lumped models*, IEEE Trans. Automatic Control, AC-21 (1976), pp. 538–541.

[31] L. M. SILVERMAN AND H. E. MEADOWS, *Equivalence and synthesis of time variable linear systems*, Proc. 4th Allerton Conf. Circuit and Systems Theory, 1966, pp. 776–784.

[32] J. C. WILLEMS, *Dissipative dynamical systems, Part II: Linear systems with quadratic supply rates*, Arch. Rational Mech. Anal., 45 (1972), pp. 352–393.

[33] ———, *Least squares stationary optimal control and the algebraic Riccati equation*, IEEE Trans. Automatic Control, AC-16 (1971), pp. 621–634.

[34] W. A. WOLOVICH, *Linear Multivariable Systems*, Springer-Verlag, New York, 1974.

[35] E. WONG, *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill, New York, 1971.

[36] W. M. WONHAM, *On a matrix Riccati equation of stochastic control*, this Journal, 6 (1968), pp. 681–697.

[37] D. C. YOULA, *On the factorization of rational matrices*, IRE Trans. Information Theory, IT-7 (1961), pp. 172–189.

[38] D. C. YOULA AND P. TISSI, *n-port synthesis via reactance extraction—Part I*, IEEE Intern. Convention Record, 14 (1966), pp. 183–205.